

Big Data transfer over Internet

Dmitry Wlasov, Vladimir Grudinin, Oleg Sadov, Andrey Shevel,
Sergey Khoruzhnikov, Arsen Kairkanov

ITMO University, St.Petersburg (Russia)



Reporter:

Andrey Shevel

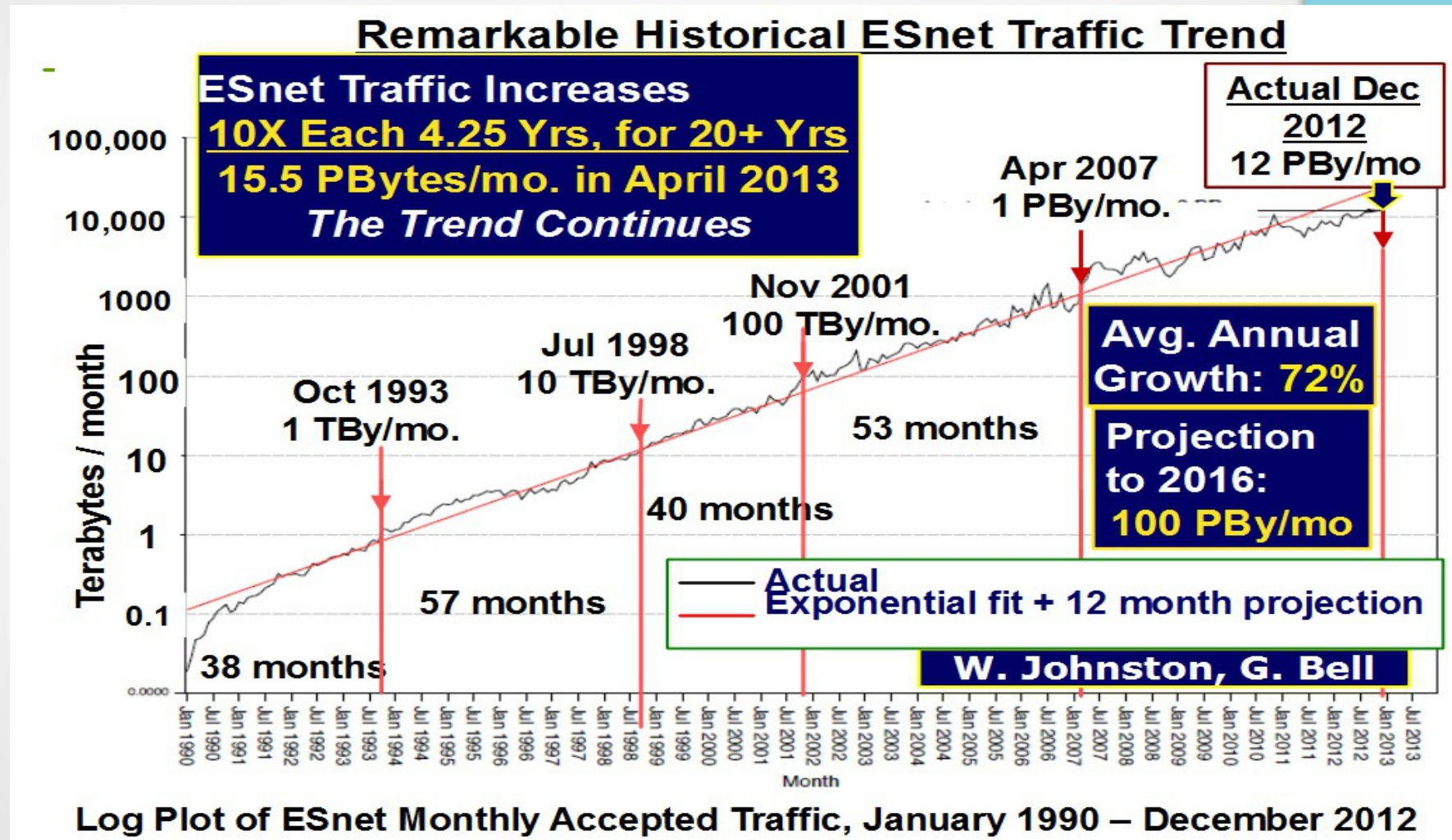
Outlook

- Sources of the Big Data.
- Architecture of the Big Data.
- Technology of the Big Data transfer.
- Our recently started research.

Scientific sources of Big Data

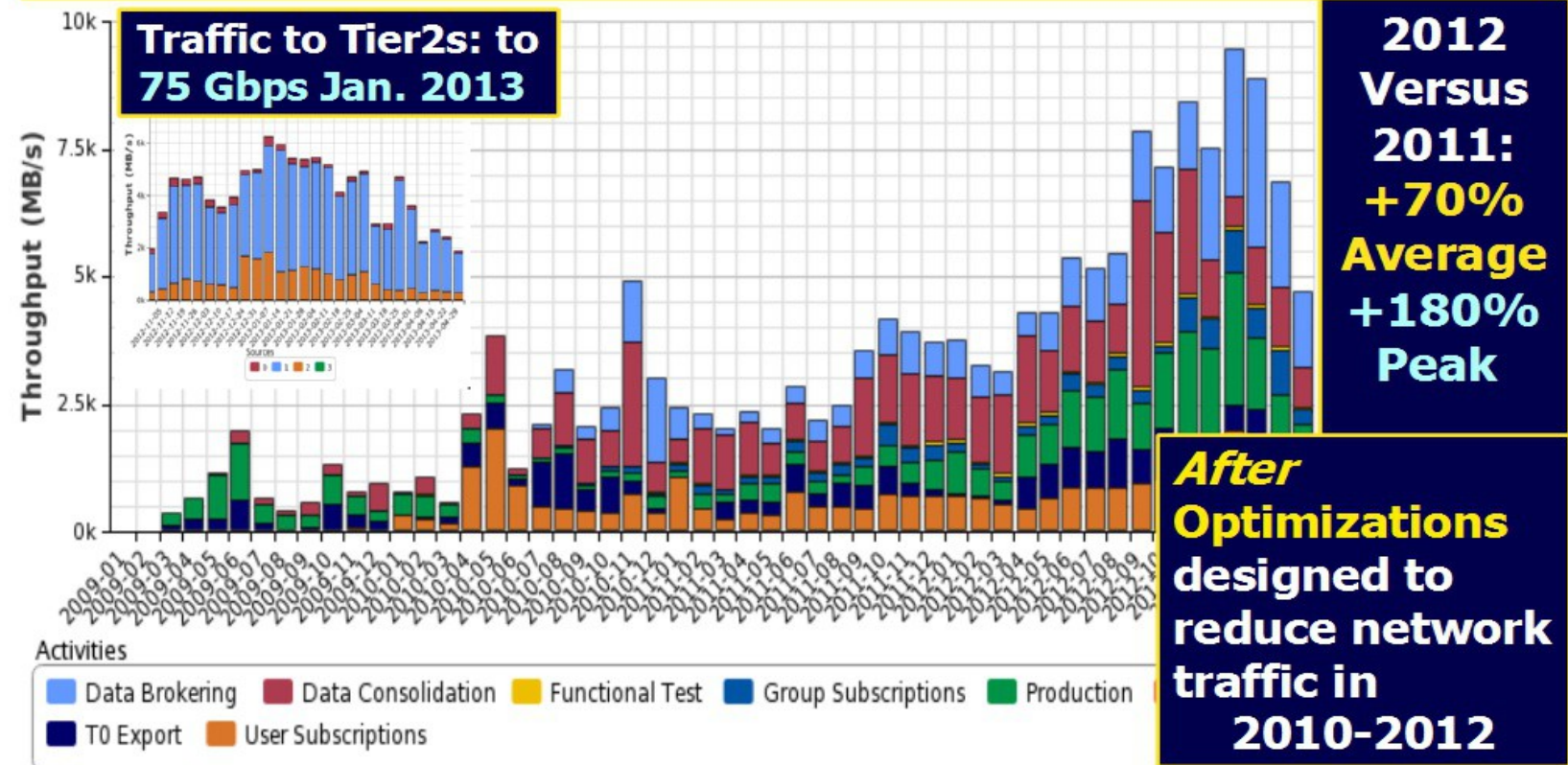
- **Scientific experimental installations**

- <http://www.lsst.org> - Large Synoptic Survey Telescope
 - **15 TB per night (may be 10 PB/year)**
- <https://www.skatelescope.org/> - Square Kilometre Array
 - **300-1500 PB/year**
- <http://www.cern.ch> — CERN
 - **~20PB/year** (FAIR ~ same)
- <http://www.iter.org> - International Thermonuclear Experimental Reactor
 - **~1 PB/year**
- <http://www.cta-observatory.org/> - CTA - The Cherenkov Telescope Array
 - **~20 PB/year**



ATLAS Data Flow: 2009- April 2013

2012-13: >50 Gbps Average, 112 Gbps Peak
171 Petabytes Transferred During 2012





Network Trends in 2012-13

100G Evolution; Optical Transmission Revolution

Increased multiplicity of 10G links in Major R&E networks:

- Internet2, ESnet, GEANT, and leading European NRENs
- **Transition to 100G next-generation core backbones: Completed in** Internet2 and Esnet in 2012; **US 100G endsites proliferating !**
- **GEANT transition to 100G: Phase 1 Completed by Mid-2013**
- **NREN 100G already appeared and spreading in Europe and Asia: e.g.** SURFnet & Budapest - CERN; Romania, Czech Rep., Hungary, China, Korea
- **100G Transatlantic (Initial trials) in 2013**
- **Proliferation of 100G network switches and high density 40G data center switches. 40G servers (Dell, Supermicro) with PCIe 3.0 bus**
- **Higher Throughput: 300G+ at SC12 – UVic, Caltech, Mich., Vanderbilt**
- **Trend towards SDN (Openflow, etc.): a Major Focus taken up by much of the global R&E network community and industry**
- **Advances in optical network technology even faster: denser QAM modulation; 400G in production (RENATER); 1 Petabit/sec on a fiber**

The move to the next generation 100G networks is well underway and accelerating; 200G, 400G production networks not far away

Which are the future demands

Paul Sheldon & Alan Tackett
Vanderbilt University



10 Year Projections

■ Requirements in 10 years? Look 10 years back...

Source: Ian Fisk and Jim Shank

Metric	<u>Tevatron</u> (2003)	<u>LHC</u> (2012)
Remote Computing Capacity	15kHS06 (DZero Estimated)	450kHS06 (CMS)
User Jobs launched per day	10k per day	200-300k jobs per day
Disk Capacity per experiment in PB	0.5PB	60PB
Data on Tape per experiment	400TB	70PB
MC Processing Capacity per month for Full Simulation	3M	300M
Data Served from <u>dCache</u> at FNAL per day	25TB per day	10PB per day
Wide Area networking from host lab	200Mb/s	20000Mb/s
Inter VO transfer volume per day	6TB (<u>DZero</u> SAM)	546TB (ATLAS)

Emerging Data Logistics Needs of the LHC Expts

17

September 5, 2013

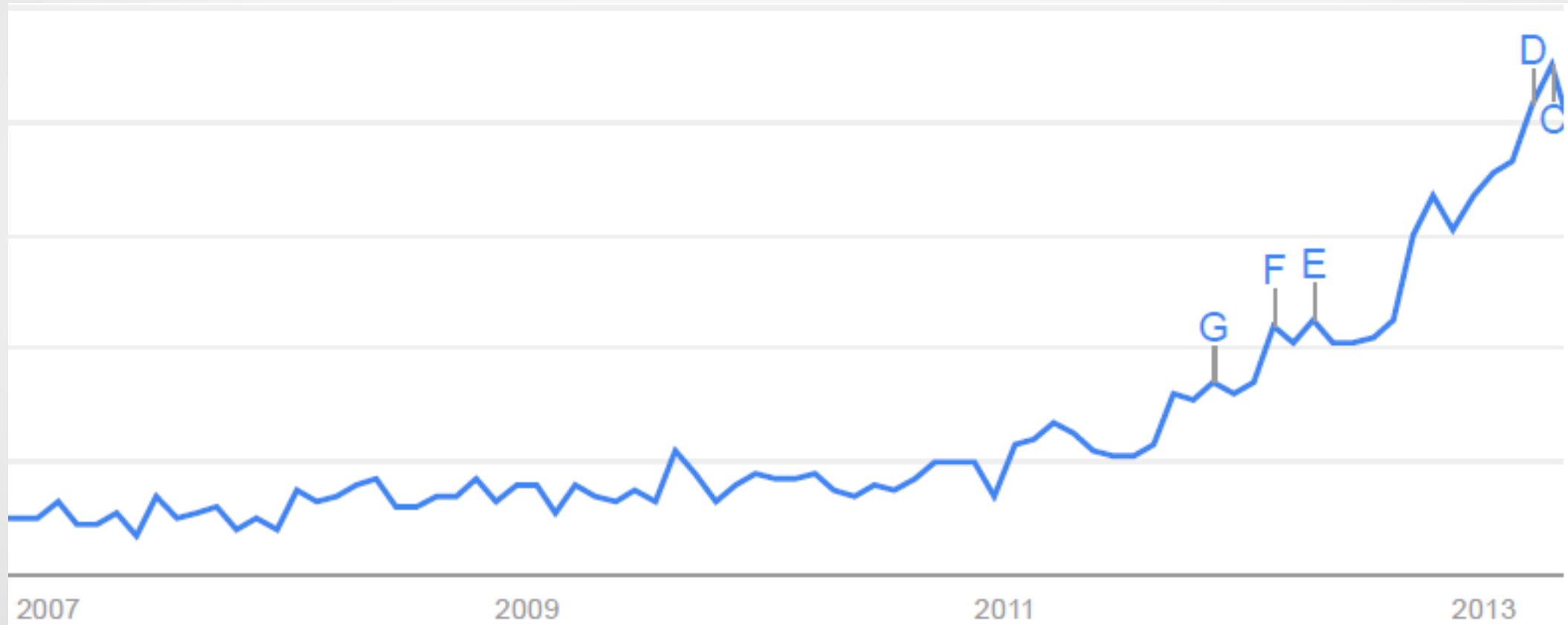
LHC Network Infrastructure

- LHC OPN — Optical Private Network
- LHC ONE — Open Networking Environment
 - the aim of LHCONE is to keep many entry points to private optical network for LHC T1/2/3 .
- ANSE — Advanced Network Services for [LHC] Experiments

Other Big Data providers and what was done to study

- **NIST** - <http://bigdatawg.nist.gov/usecases.php> - many examples and use cases
 - Any communication company.
 - Video streams from tens of thousands of watching cameras.
 - Data Preservation.

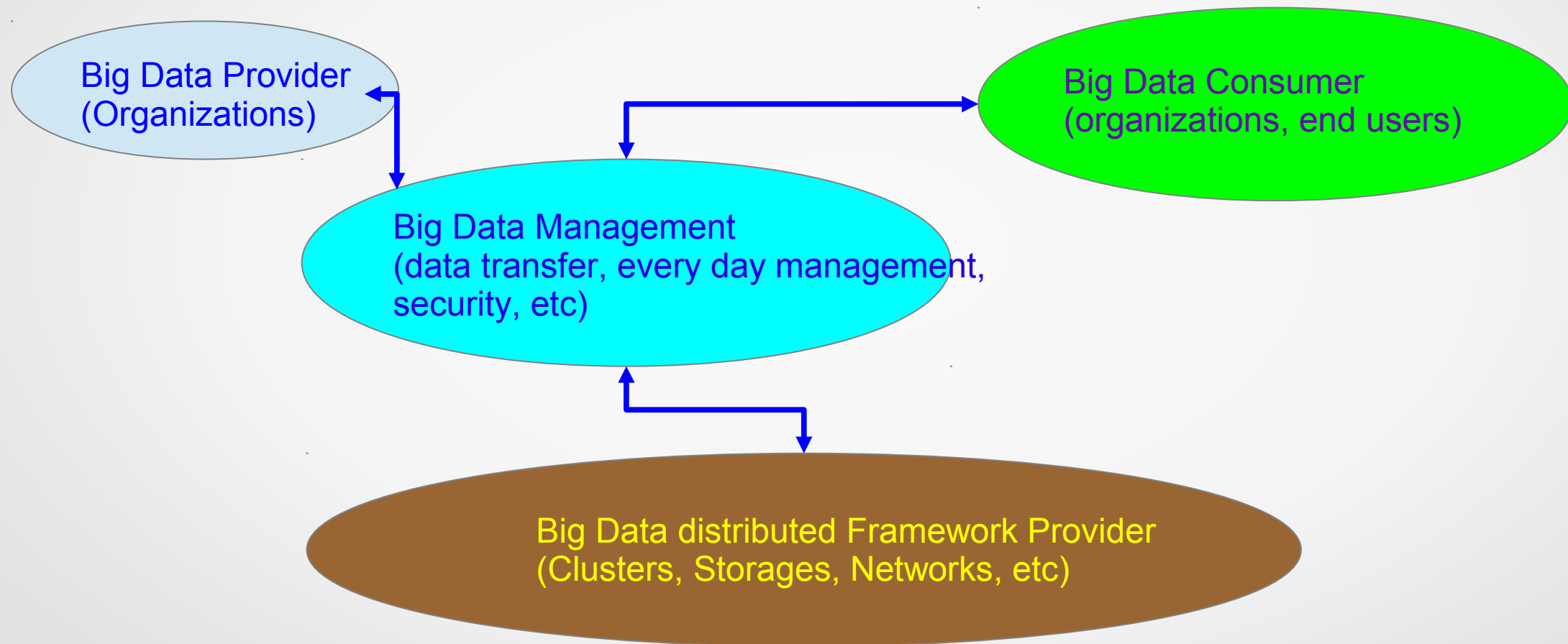
Google trend for «the big data»



Big Data

- Big Data have the set of features:
 - Data flow Velocity
 - Data Volume
 - Data Variety

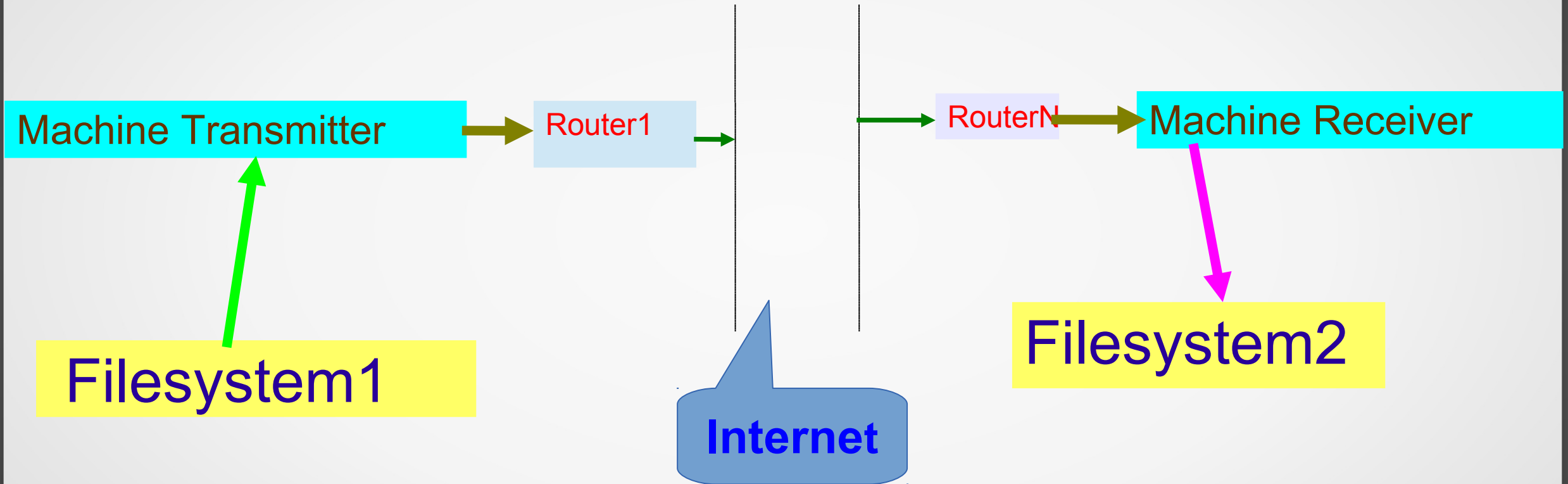
Big Data architecture



Peculiarity of the Big Data transfer

- Big Data transfer might consume many hours or days.
- The situation in channel might be changed: RTT, % of lost network packages, data link bandwidth).
- Finally, it might occurred the interruption (hours?, days?) in operation of data link .
- Obviously it is useful to have access to two or more independent data links.

Process of the Data transfer



Research topic at ITMO University: the transfer of Big Data

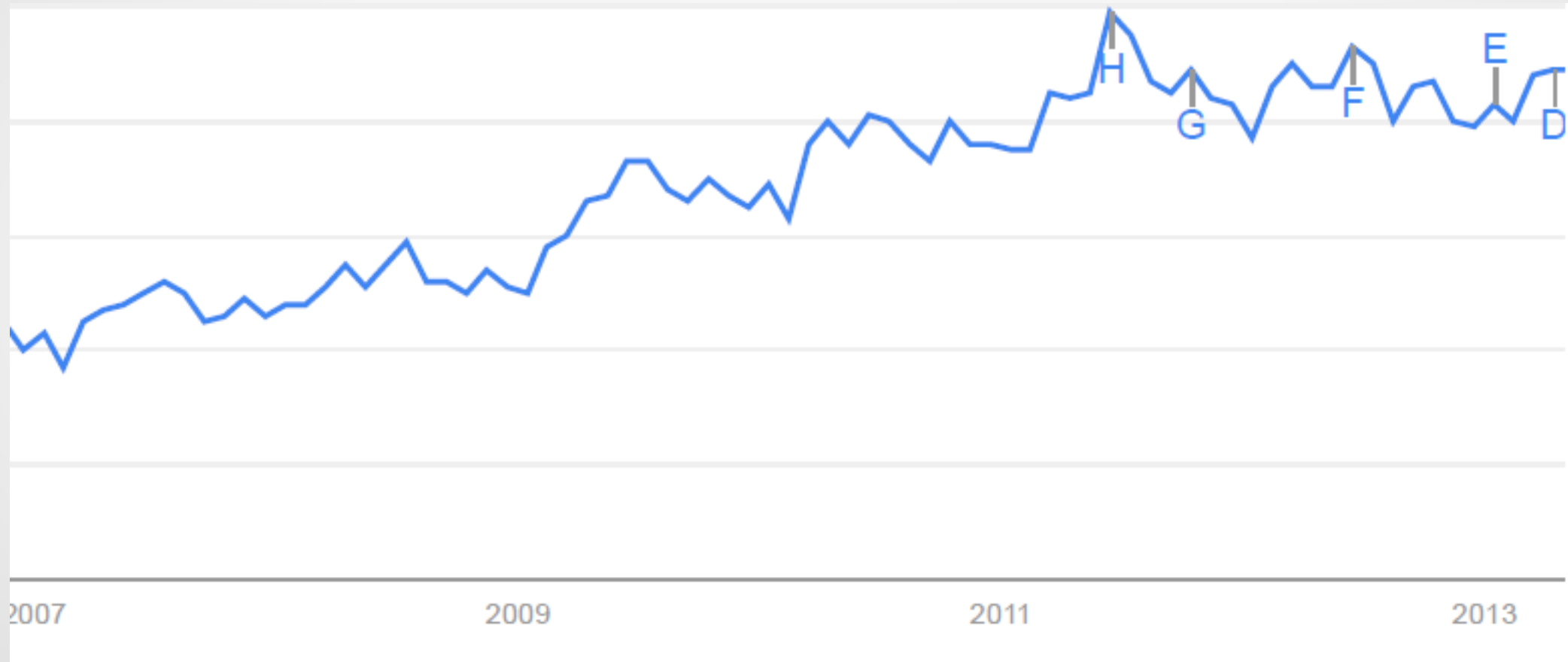
In laboratory of network technologies <http://sdn.ifmo.ru/> at ITMO university <http://www.ifmo.ru/> the new research «Big Data transfer over Internet» has been formed .

- The special testbed (100 TB of disk storage + server 96 GB of main memory under OS RedHat/ScientificLinux on each side).
 - Comparative study of the existing tools of the data transfer (testing and measurements).
 - To use the testbed as instrument to compare various tools (tracking for the measurements + results).

Planned measurements

- Long distant sites with existing data links (not only most advanced links).
- The idea is to use more than one data link in parallel.
- Recently we obtained some experience with Software Defined Networks (SDN) and now we plan to use that approach in the Big Data transfer.
- We are starting to collaborate with GENI (<http://www.geni.net/>)

Google trend for SDN



Technology peculiarities with Big Data transfer

- Still main protocols — stack of TCP/IP.
 - Number of network parameters in Linux (around 1/2 thousand).
 - Important parameters: e.g. size of block, size of TCP Window.
 - Main method to decrease the transfer time (even over one data link) is using the multi-stream data transfer.

Testing on the first stage (program tools)

- GridFTP - <http://www.globus.org/toolkit/data/gridftp/>
- BBFTP - <http://doc.in2p3.fr/bbftp/>
- FDT - <http://monalisa.cern.ch/FDT/>
- FTS3 - <http://fts3-service.web.cern.ch/>
- Also technology components to watch the data links status, e.g. perfSONAR.

Partners (information exchange)

- On Big Data

Laboratory of Information Technology (LIT) <http://lit.jinr.ru/index.php?lang=lat>

@ Joint Institute for Nuclear Research (JINR.ru)

- On Software Defined Networks

- The center of Application Research Center for Computer Networks at Moscow University <http://arccn.ru/>

Ideas to compare the data transfer tools

- Availability.
- API.
- Performance.
- Reliability.
- Operation tracking.
- Ability to predict the time to transfer the data on the base of existing tracking records.
- Required resources: memory, CPU time, etc.
- Others.

Main goals

- Combining the developed contemporary components and methods with ideas, developments, experience achieve maximum speed for Big Data transfer on existing links.
- To create the testbed which would be used as place where researchers might compare theirs (new) tools for data transfer with previously recorded measurement results.
- To suggest the collaboration to ... (suggestions?)
- To invite students ...(suggestions?)

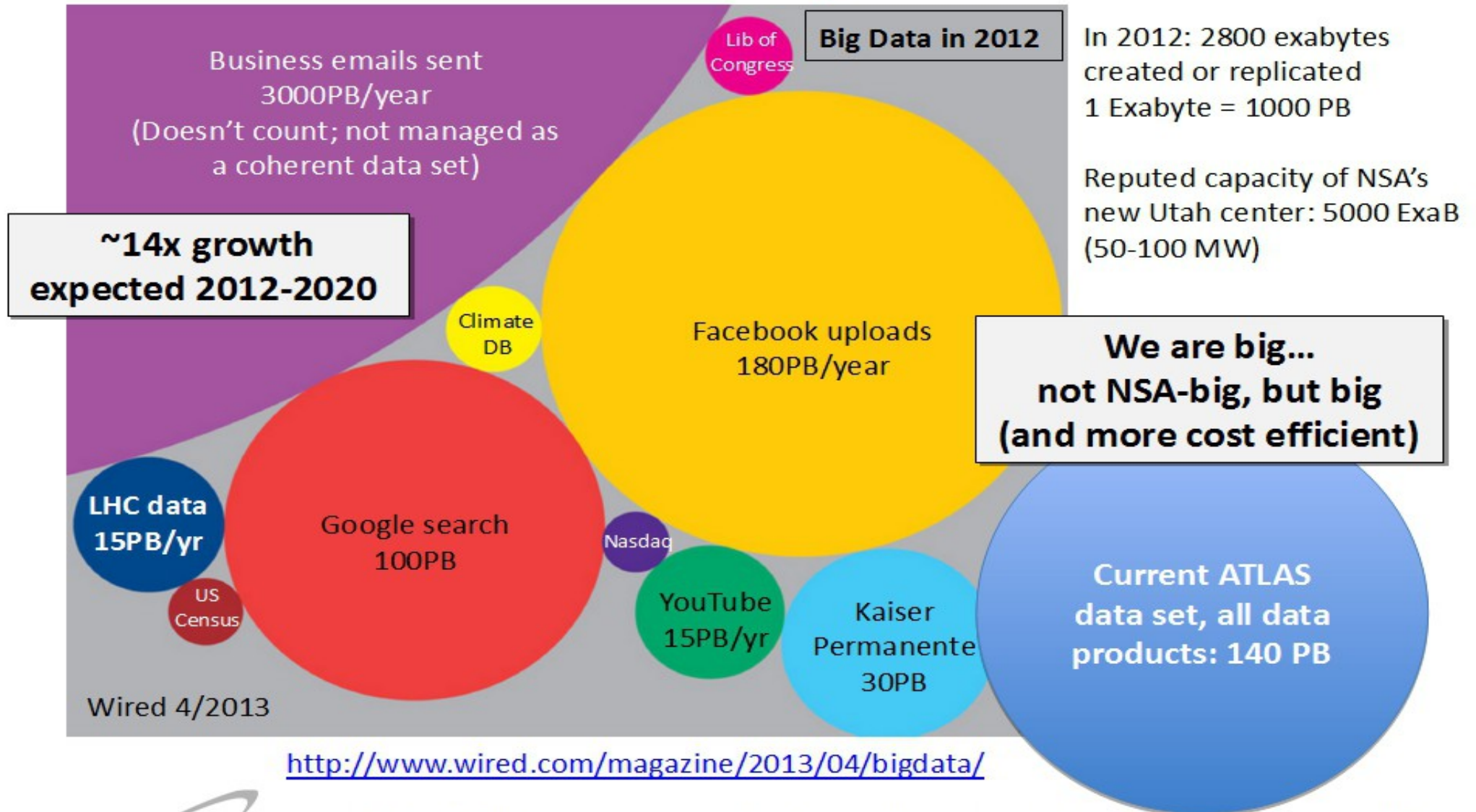
What was done until now

- There were deployed
 - two servers under Scientific Linux 6.5.
 - Openstack with appropriate set of Virtual Machines to test a number of mentioned utilities.
 - PerfSonar
- Our research laboratory site is <http://sdn.ifmo.ru>

Questions ?

Data Management

Where is LHC in Big Data Terms?



Decimal

Value		Metric
1000	KB	kilobyte
1000^2	MB	megabyte
1000^3	GB	gigabyte
1000^4	TB	terabyte
1000^5	PB	petabyte
1000^6	EB	exabyte
1000^7	ZB	zettabyte
1000^8	YB	yottabyte